The Artificial Evolution of Cooperation

Nicolas Meuleau[1] and Claude Lattaud[2]

[1] CEMAGREF, Artificial Intelligence lab,
Parc Tourvoie, BP 121, F-92185 Antony Cedex, France.
e-mail : nicolas.meuleau@elan.cemagref.fr
[2] Université Paris 5, LIAP 5,
45, rue des Saints Pères, F-75006 Paris, France.
e-mail : latc@math-info.univ-paris5.fr

Abstract. We propose here a new approach to study co-evolution and we apply it to the well-known iterated prisoner's dilemma. The originality of our work is that it uses a simplified version of the game, and thus, restrict the search space of evolutionary dynamics. This allows to have a look at the totality of the search space in permanence, and so, a complete understanding of the phenomenon of co-evolution in process. The paper includes a little game-theoretic introduction to iterated prisoner's dilemma, a survey of previous works on evolution in this game and the exposition of the questions that were still asked to us. We describe then our special approach to the problem, using populations larger than the search space, or even infinite. The experimental results that we present complete the actual knowledge of iterated prisoner's dilemma.

1 Introduction

In a symmetric two-players game, each one secretly chooses a decision among the N available. The $N \times N$-matrix $\nu = [\nu_{ij}]$ that defines the game, gives the utility earned by each player consequently to these decisions : ν_{ij} is the result of the player who chooses decision i while its opponent chooses decision j.

In the Prisoner's Dilemma (PD), there are two possible decisions (i.e. $N_{\text{PD}} = 2$), number 1 being interpreted by "Defect" (D) and number 2 by "Cooperate"(C). The utility matrix is given by the following table :

player 1 \ player 2	D	C
D	$P\,/\,P$	$T\,/\,S$
C	$S\,/\,T$	$R\,/\,R$

(result of player 1 / result of player 2)

i.e. :

$$\nu^{\text{PD}} = \begin{bmatrix} P & T \\ S & R \end{bmatrix},$$

where $T > R > P > S$. We see that for a very wide variety of parameters, PD is a non-zero-sum game : a reward for a player does not enforce a cost for its “opponent”.

Player 1 may lead the following reasoning :

– “If player 2 cooperates, then it is preferable to defect, because the *Temptation* of defection T is greater than the *Reward* for mutual cooperation R.”
– “If player 2 defects, then it is again better to defect, because the *Penalty* for mutual defection P is greater than the *Sucker*'s punishment S.”

This may be represented by drawing two arrows on the table in the following way :

player 1 \ player 2	D	C
D	$P\,/\,P$	$T\,/\,S$
	↑	↑
C	$S\,/\,T$	$R\,/\,R$

The same reasoning step can be applied to player 2, what leads to the drawing of two symmetric arrows :

player 1 \ player 2	D		C
D	N $P\,/\,P$	←	$T\,/\,S$
	↑		↑
C	$S\,/\,T$	←	$R\,/\,R$

So, if both player follow the rational reasoning developed above, they both conclude that defection is the best alternative whatever the opponent's choice. It results in the situation (D, D), convergence point of the arrows. This fact is formalised by saying that (D, D) is a Nash equilibrium of the game:

Definition 1. (i, j) is a Nash equilibrium of game ν iff. :

$$\nu_{ij} \geq \nu_{i'j} \quad \forall i' \neq i, \quad \text{and} \quad \nu_{ji} \geq \nu_{j'i} \quad \forall j' \neq j.$$

That is why this situation is marked with a N on the table.

So the defection is the natural choice of two rational, but not speculator at all, agents. However this choice seems somehow non-optimal because both players could obtain a greater gain by cooperating ($R > P$). We will say that (C, C) strictly dominates (D, D) :

Definition 2. (i, j) strictly dominates (i', j') in the game ν iff. :

$$\nu_{ij} > \nu_{i'j'}, \quad \text{and} \quad \nu_{ji} > \nu_{j'i'}$$

All the dilemma lies in the fact that the Nash equilibrium of the game is strictly dominated by a diametrically opposed situation [3] . For this reason, PD is called a strong cooperation problem.

In PD, the players are supposed to meet only once [4] . In the Iterated Prisoner's Dilemma (IPD), the two same players are supposed to meet and play PD several times successively. They are allowed to have their decision at the n^{th} meeting depend on all the past interactions with the same opponent, i.e. the results of the $n - 1$ previous meetings. The aim of each one is to maximise the summation over all iterations of his rewards [5].

The constraint $T + S < 2R$ is imposed on the utility values. Its role is to guarantee that two players that always cooperate will do better than two players that have agreed to play (C, D) and (D, C) alternatively.

Theoretically, IPD is a two-player symmetric game exactly as PD was, we note uIPD its utility matrix. In IPD, a decision is the choice of a strategy, i.e. of a mapping from the set of all possible sequences of results – the "past histories" – into the set D, C. As we show below, there is a huge combinatorial explosion of the decision space of IPD.

We call "length of the interactions" or "number of iterations" L, the number of times that two same players play PD together. The number of different possible histories $x(L)$ may be iterativelly calculated with the following formulas :

$$x(1) = 1, \quad \text{and} \quad x(L) = 4^{(L-1)} + x(L-1) \quad \forall L > 1.$$

The first one comes from the fact that, if $L = 1$, then IPD equals PD and there is only one possible history, noted $\emptyset$: "nothing has been played, the game begins". After the first iteration, the game is already finished.

The second formula uses the fact that the result of each iteration takes its value in the set T, R, P, S. There exists so $4^{(L-1)}$ different possible histories of exactly $L - 1$ iterations. The second term of r.h.s. comes from the fact that a situation where less than $L - 1$ iterations have been played is a possible history in a game of length L.

From the definition of a strategy as a mapping from a set of size $x(L)$ into a 2-elements set, it results that the number of decisions of IPD is : $N_{\text{IPD}} = 2^{x(L)}$.

[3] in the sense that every player must change his decision to go from the dominated situation to the dominating one.

[4] this is equivalent to suppose that the players randomly change of opponent at each turn and never know who they are playing with (anonymous game), or that the players forget all their past experiences between two iterations, or also that they are myopic and cannot anticipate farther than one time-step.

[5] sometimes, it is the discounted cumulated reward (cf. Axelrod 1984) that is to maximise. In regard to the experimental aspect of our work, the non-discounted reward is easier to handle. Moreover, the robustness of our results seems to indicate that the use of discounted rewards would not affect a lot the behaviour of the model.

So it increases exponentially relatively to the exponential of L. For instance, if $L = 4$, two players play only four times together but $N_{\rm IPD} \simeq 4\,10^{25}$. This number is not computable and so, even with this very short length of interactions, the combinatorial explosion of the search space is getting too much for us.

With the repetition of non-anonymous interactions, a lot of high-level concepts as benevolence, malevolence, susceptibility and indulgence are introduced into the game. In a lot of fields as varied as sociology, psychology, international politic and biology, IPD has proven to be a very useful and powerful model (cf. Axelrod 1984 ; Baefsky and Berger 1974 ; Bethlehem 1975 ; Riker and Brams 1973 ; Schelling 1981). Lately works on IPD are almost all devoted to the emergence of cooperation in an evolving population of players. Since his leading works and the publication of his book "The Evolution of Cooperation" (1984), R. Axelrod's name is almost always associated with this field of study . In the following section, we give a survey of his results and the most relevant of associated works.

2 Previous Works

The first topic presented in R. Axelrod 's book "The Evolution of Cooperation" is the result of computer tournaments opposing several strategies for IPD proposed by researchers from different spheres. As it is well known, the winner of these tournaments was the strategy "Tit for Tat" (TFT), that may be defined in the following way : cooperate at the first iteration and after, always do whatever your opponent did on the previous iteration. The author attributes this success to following special properties of TFT :

benevolence : never to be the first to defect,
susceptibility : to punish a defection of the opponent by another defection,
indulgence : to forgive the opponent after having punished him.

The next step of Axelrod studies is the simulation of what should have happened if he had continued to organise tournaments. Assuming that :

- no new strategy should have been invented and introduced into the tournaments,
- the representation of a strategy in a tournament (i.e. the percentage of people using it) is proportional to its average score at the previous tournament[6],

he leads computer simulations of tournaments. These experiments show that TFT would have continued to dominate the tournaments, in percentage of representation and in score, as long as the two hypothesis above are still respected.

After having verified the superiority of TFT in the artificial environment constituted by the set of strategies proposed by scientists, Axelrod leads a theoretical study of "evolutionary stability" to explain this result. For this purpose, he defines the notion of collective stability in the following way :

[6] in a way very similar to Genetic Algorithms'selection operator.

Definition 3. The decision i for game ν is collectively stable iff. :

$$\nu_{ii} \geq \nu_{ji} \quad \forall j \neq i$$

So, a strategy i is collectively stable if (i, i) is a Nash equilibrium (according to definition 1) of the game. Then we say that i is in Nash equilibrium with itself.

The reasoning laid by Axelrod is the following : if the population is uniquely constituted by a single "indigenous" strategy i, then, to be able to survive, an invading strategy j must do a score strictly greater than the natives. As the invader and most of the indigenous play only against native i, because of their great numerical superiority, j will survive if $\nu_{ji} > \nu_{ii}$. Thus strategy i is evolutionary stable if this inequality never holds.

Using this definition, he shows that TFT is collectively stable in IPD. Then he states that collective stability is evolutionary stability and so, TFT is evolutionary stable. This work is closely related to J. Maynard-Smith's studies (1975), although Axelrod's collective stability is less restrictive than Maynard-Smith's evolutionary stability.

A doubt is cast on the utility of these results results by Boyd and Loberbaum (1987). The critic is that it is sufficient that the invading strategy j realises $\nu_{ji} \geq \nu_{ii}$ to be able to survive in an homogenous population of i. As the authors notice, this distinction is important in IPD because a lot of couple (i, j) verify $\nu_{ji}^{\mathrm{IPD}} = \nu_{ii}^{\mathrm{IPD}}$. It is to be noted that, after having argued that the collective stability of TFT is not sufficient to explain the emergence of cooperation, the authors propose another definition of evolutionary stability, more restrictive than Axelrod 's collective stability. Then they show that no deterministic strategy may be evolutionary stable in IPD according to their definition.

It seems to us that this criticism is very well founded. It is particularly pertinent in the case of evolutionary algorithms as Genetic Algorithms (GA) and Evolution Strategies (ES). See (Bäck and Schwefel 1993 ; Hoffmeister and Bäck 1991) for an overview of evolutionary algorithms, (Holland 1975 ; Goldberg 1989 ; Bäck and Hoffmeister 1991 ; Bäck 1992) for GA, and (Rechenberg 1973 ; Schwefel 1977 ; Schwefel 1981 ; Herdy 1991) for ES.

Imagine an evolutionary algorithm (GA or ES) when, at some time, the population is uniquely constituted with TFT. In appearance, it expresses by a population of agent that always cooperate. Evolutionary algorithms working in a somehow blind mode, random mutations may always happen. For instance, an individual playing the strategy "Always Cooperate" (ALLC) may appear. Externally, this modification is not seenable since everybody continues to cooperate. So, the ALLC mutant realises the same score as TFT indigenous, and for this reason, he has the same probability to survive selection. Thus, it is theoretically possible that little by little the TFT population becomes invaded by ALLC, creating so a favourable ground for the later appearance of "Always Defect" (ALLD) individuals. We do not know what will then happen, it seems that the population must tend to an equilibrium point, or enter a cyclic dynamic. In all cases, it is highly improbable that the population still be composed only of TFT after a few iterations of the algorithm. So TFT is not stable under the action of GA and ES [7].

A definition of stability consistent with evolutionary algorithms is the following :

Definition 4. The decision i for game ν is EA-stable iff. :

$$\nu_{ii} > \nu_{ji} \quad \forall j \neq i$$

So, a strategy is EA-stable if it is in strict Nash equilibrium with itself, according to the definition :

Definition 5. (i, j) is a strict Nash equilibrium of game ν iff. :

$$\nu_{ij} > \nu_{i'j} \quad \forall i' \neq i, \quad \text{and} \quad \nu_{ji} > \nu_{j'i} \quad \forall j' \neq j.$$

It is be noted that our EA-stability is more restrictive than Boyd and Loberbaum's evolutionary stability (but this rough definition is enough to illustrate our purpose). So, we can deduce of their results that *no deterministic strategy is EA-stable in IPD.*

Although Axelrod's theoretical results may be put in the balance, he also proposes experimental results to support the thesis that cooperation based on benevolence, susceptibility and indulgence is a natural phenomenon. In (Axelrod 1987), he presents results of simulations where a population of 20 strategies for IPD is submitted to the action of a GA. Each one satisfies :

Hypothesis M_3**.** The choice of every decision only depends on the three last results with the same opponent (3 steps of memory),

and is encoded on a linear chromosome of length 70. This experiment concludes with the emergence of cooperation, due to the dominance of TFT and TFT-like individuals.

A lot of other studies have been laid to experiment the evolution of a population of IPD-strategies under the action of evolutionary algorithms. The main differences between them is the way to encode strategies, what directly determines the nature of the search space (Fogel 1993 ; Gacögne 1994 ; Lindgren and Nordal 1994 ; Mühlenbein 1992). Most often, the authors succeed in having cooperation emerge. It seems to us that, because of its simplicity, Axelrod's result is still the most convincing. However, we think that the search space used in this study is too large beside of the population size to allow a deep understanding of the mechanisms in process.

[7] all this reasoning may be summarised in the following way : Axelrod's collective stability is consistent with a model of evolution where there is a cost imposed on mutation. This is not the case of GA and ES, where the calculus of the fitness of an individual does not take into account the fact that it is a mutant or an indigenous.

3 Our Approach

As we have seen in section 2, the decision space of IPD is totally out of the range of actual computers. We are concerned with the following of its consequences : a simulated population of strategy for IPD may only cover an insignificant part of the complete search space.This is a issue very important to keep in mind when we approach IPD and evolution.

Even if they restrict their study to a sub-set of the complete strategy space, previous works on evolution and IPD always use search spaces too large for being totally explored. For instance, Axelrod's sub-set of strategies satisfying hypothesis M_3 is of size $2^{70} \simeq 10^{21}$, and he uses a population of only 20 individuals. Moreover, the size of the search space seems to be one of the main limitations to the efficiency of evolutionary algorithms. The following appealing remark is made about the validity of D. E. Goldberg's Schema Theorem by Bäck and Schwefel (1993) :

> "*. . . finite populations often do not contain all instances of a specific schema. Observed schema fitnesses thus might quite mislead the search process.*"

In order to be able to lead reliable simulations with population sizes not too low beside of the search space, we have chosen to limit ourselves to the set of the 32 strategies that satisfy :

Hypothesis M_1. the choice of every play only depends on the result of the last iteration played with the same opponent (1 step of memory).

It appears to us that this is the easiest way to arbitrarily restrict the search space to a very low sub-set of the set of all strategies, while preserving the symmetry of the original space. Moreover, most of the atypical strategies as TFT, ALLC and ALLD satisfy M_1. An arbitrary numbering for the strategies satisfying M_1 is proposed in table 1. It will be used all over the end of this paper.

When we limit ourselves to the strategies satisfying M_1, IPD becomes a 32-decisions game. Its 32×32 utility-matrix ν^{IPD} may then be easily calculated by simulating, on any computer, all the possible confrontations of two strategies, for a fixed length of interaction L. We can then verify that, as it is announced in section 2, there is no strategy that is EA-stable, whatever the value of L.

However, there may exist mixed situations (i.e. situations where different strategies cohabit) that constitute equilibrium points of the system. The analytical determination of them is possible by a solving 32-unknowns system, after having defined the dynamic of the system (cf. section 4). However we prefer to lead an experimental determination of them. Our idea is to try different evolution schemes on a population of strategies satisfying M_1.

Our work could seem to be "another" IPD-based simulation of evolution. However, the originality of our work lies in the fact that we can use population sizes greater than the search space, and then lead simulations where the observed behaviour of the algorithm is not too far from its expected behaviour. In the limit, we represent the population by a set $\{q_1, q_2, \ldots q_{32}\} \in [0;1]^{32}$ such that :

$$\sum_{i=1}^{32} q_i = 1$$

(i.e. belonging to the simplex S_{32} of $\mathbb{R}^{32}$) where q_i represents the percentage of individual using strategy i in the population.

With this continuous approach to the population, we associate a deterministic handling of chance : all calculus are made by using the expected values of variables with respect to calculated probabilities, instead of by simulating random drawings, as it is the case in classical GA and ES. So, as a consequence of the weak law of large numbers, the continuous approach to the population simulates the theoretical behaviour of the model with an infinite population. It allows a very interesting look at the evolution of the population that would not have been possible in greater search spaces.

To present our model, we develop in section 4 a discussion about the question : "what is an evolutionary dynamic ?".

4 Artificial Evolution

The actual model of natural evolution is the following :

1. each individual metabolism is determined by its genetic code, the DNA, that is proper to him and present in all of its cells.
2. DNA is the support of heredity, it is transmitted from parents to children, on the condition of some random recombinations and mutations.
3. An individual metabolism determines its external characters, the environment selects the individuals possessing the characters the most favourable for survival and reproduction. This phenomenon of natural selection modifies the statistical distribution of genetic codes from generations to generations.

Evolutionary algorithms are based on the idea that there exist an intrinsic "optimising principle" in this scheme, and they try to reproduce it in purpose of optimisation. But, because they only modelise a part of the whole process, the mechanism processed is quite different.

GA basic principle is to replace the problem of optimising a numerical function f of some discrete set X, by the problem optimising a numerical function g of a certain set C of "codes" of X elements. The *Building Block Hypothesis* (Holland 1975 ; Goldberg 1989) states that we wait from the code to satisfy some special properties with respect to the objective function. In particular, this is needed that individuals obtained by recombination (at the code level) of good individuals, are also good individuals. Somehow, things may abstracted the following way : by defining notions of neighbourhood and distances, the recombination operator induces a topology on the code space C ; the function g is then supposed to have some properties close to continuity with respect to this topology.

The *Strong Causality Principle* , one of the two fundamental principles of ES (Hoffmeister and Bäck 1991 ; Herdy 1991), explicitly stipulates that a small variation of the phenotype induces a small variation of the objective (fitness) function. So, ESs directly use a property of continuity of the objective function.The need of a certain continuity of the fitness function in evolutionary mechanisms also appears in (Feistel and Ebeling 1989), that greatly inspired us for building our model. As these authors do, we modelise the evolutionary process with two operators :

mutation : a mechanism of random transformation of an individual into one of his neighbours, with respect to a certain topology. With infinite population a diffusion operator is used.

selection : a mechanism that determines the individuals able to survive to the change of generation, deterministically or stochastically, and in a way that favours the performance.

We tried several alternatives for the choice of the mutation topology and the selection operator. With these choices, we wanted to build a system not to far from GA and ES. However we do not argue that our scheme is an exact reproduction of all these algorithms. This particularly true in the case of GA, because of the absence of explicit recombination in our model.

Even if it does not always constitute a good model of evolutionary algorithms, we argue that our system is a not so bad model of evolution. In all cases it is very close to Feistel and Ebeling's one. Moreover, the surprising robustness of our observations (cf. section 5), allows to conjecture that the behaviour of the system would not have change if we had used an explicit recombinative operator.

There also exist an other point where our model is closer to Feistel and Ebeling's one than to evolutionary algorithms. In GA and ES, the fitness of an individual is constant and deterministically determined by the objective function f. In our model, the fitness f_i of strategy i depends on all the population, and so, it varies through time. Defining q_i as the percentage of individual using strategy i in the population, then we use :

$$f_i = \sum_{j=1}^{32} q_i \, \nu_{ij}^{\mathrm{IPD}}.$$

The underlying hypothesis is explained in the next paragraph. We want to quote now that the use of a formula of type $f_i = f(q_1, q_2, \ldots q_N)$ to replace in Feistel and Ebelling's model and in ours, the less general formula $f_i =$ a constant $= f(x_i)$ of evolutionary algorithms, could be used to differentiate evolutionary computation as a particular field of a more general artificial evolution domain. The artificial evolution approach allows to really modelise evolving interactions, and so, is much closer to the process of natural selection than evolutionary computation models. Previous works on IPD and evolution like (Axelrod 1987 ; Fogel 1993 ; Lindgren and Nordal 1994) are artificial evolution models in the sense defined here. This also the case of artificial ecology models as (Werner and Dyer 1991).

By choosing

$$f_i = \sum_{j=1}^{32} q_i \, \nu_{ij}^{\mathrm{IPD}}.$$

we implicitly made the hypothesis of no spatial distribution of the population. With a finite population, it formulates as :

Hypothesis NS. At each generation, every individual meets every other one once, its fitness is its average score over these meetings (no spatialization).

With infinite population, it formulates as :

Hypothesis NS'. At each generation, every individual meets a big number of its colleagues and its opponents are randomly drawn according to the probabilities (q_i).

5 Experimental Protocol

We give below an exhaustive list of all the parameters of the model. Readers desiring more information are invited to contact us.

IPD parameters : i.e. the utilities P, T, S and R, and the length of interactions L. They are used when the simulation begins, to calculate ν^{IPD} by simulating all possible confrontations.

Population size S : that may be infinite, as explained in section 3. It is interesting to see how large must the population be, to have the observed evolution of the system correspond with the expected evolution. We may so verify the accuracy of the quote from Bäck and Schwefel (1993) in section 3

Topology of mutation : we tried two alternatives :

- trivial topology : each strategy is the neighbour of every strategy ;
- structured topology : a neighbour of a given strategy i is obtained by flipping one decision in the rule representation as used in table 1. Thus every strategy has exactly 5 neighbours.

Selection operator : proportional or elitist :

- proportional selection is a GA selection operator. The probability of selection of strategy i is equal to :

  $$\frac{q_i \, f_i}{\sum_{j=1}^{32} q_j \, f_j},$$

  where f_i is the fitness of strategy i and q_i is the percentage of individuals using strategy i in the population.
- elitist selection is the selection operator used in $\lambda + \mu$-ES (Bäck and Schwefel 1993 ; Hoffmeister and Bäck 1991). Mutation creates new individuals and increases the population size. The selection brings it back to its initial value by suppressing the worst individuals first.

Temperature θ : the parameter that determines the magnitude of genetic mixing at each generation. With proportional selection, it is a probability of mutation, and with the elitist selection, it is to the ratio μ/λ.

Initial population : uniformly distributed or composed of a single strategy that may be chosen among the 32 possible ones.

With some special sets of value attached to different parameters, our system correspond to existing evolutionary algorithms :

- with elitist selection and trivial topology, our system is a $\lambda + \mu$-ES with no recombination and where $\lambda = S$ and $\mu = S\,\theta$.
- with proportional selection and structured topology, it is a GA with no crossover and a special mutation operator that works at the scale of individuals. The probability for an individual to mute is $p_m = \theta$.

The real-time needed to run a simulation is very short. The limit behaviour of the system (equilibrium or cycle) is always obtained in less than one minute on a good PC. So, we could try a very large variety of parameters and totally explore the model. As a consequence, an exhaustive description of the totality of our observations is impossible in a small paper like this one. That is why we have chosen to present our results in a descriptive and non-rigorous form. Some representative graphics are added to illustrate our purpose, they are gathered at the end of the paper.

6 Experimental results

This section is divided into two parts. The first describes the behaviour of the model with infinite population. Of course, this behaviour may be obtained with finite but large populations. We develop this point in the second part, that constitutes a description of the influence of all the parameters listed in the previous section.

To describe our observations with infinite population, we also put apart the case where elitist reproduction and structured topology are used together. As a matter of fact, *excepted for this special case, the population always converges to an equilibrium.* Moreover, we verify that the composition of the initial population has no influence on the state of the system at equilibrium. That seems to indicate that *this is the only stable equilibrium of the model.*

So, with proportional selection or/and trivial topology, there exists a unique equilibrium of the model dynamics. The "quality" of the equilibrium is well measured by the average score of individuals over the population when it is reached. Figure 1 represents the evolution of this data as a function of the temperature θ, with proportional and elitist selection, and trivial topology of mutation. This is to be noted that, although, they have been drawn with some special values for the parameters, the two curves have a high level of generality and similar results are obtained with other sets of parameters.

In fig.1 we verify that *the average score at equilibrium tends to 100% of cooperation when θ tends to zero* (no mutation at all). With proportional selection, this value is reached in zero, but nowhere else. In the case of elitist selection, there is a discontinuity in zero, where the score at equilibrium is the one of a uniformly distributed population. This is due to the fact that the $\lambda + \mu$-ES model implemented with elitist selection, does not affect the population in any way when $\mu = \theta\, S = 0$. Thus the population at equilibrium is the same as the initial population, i.e. uniformly distributed.

We also see in fig.1 that the equilibrium quality rapidly decreases when the temperature raises. In the case of proportional selection, we tend to a random game when the temperature tends to 1. This is not surprising since the GA model behaves almost as a random algorithm when $p_m = 1$. With elitist selection, there is a fall from 100% of cooperation to 100% of defection around $\mu = 2\,\lambda$. This is a typical instance of the brutal behaviour of the ES model due to its extremist way to carry on selection with the max operator.

We are now interested in the composition of the population at equilibrium. Figure 2 gives two very representative examples with low temperature and so, high level of cooperation at equilibrium. In these barcharts, it appears that strategy 18 is always dominating at equilibrium. Other experiments show that, *each time that cooperation emerges, 18 eventually dominates*. Referring to table 1, we may explicit the behaviour of 18 in the following way : always cooperate until the first defection of the opponent, and then, always defect until the end of the game. Thus, it is an ALLC that turns to ALLD at the first defection of the opponent. So it possesses as TFT the properties of benevolence and susceptibility (cf. section 2), but unlike TFT, it has no indulgence at all. For this reason, and following J. Maynard-Smith's inspiration, we call it "the Retaliator" (RET).

To understand the role of RET in the emergence of cooperation, we must look at the evolution of the population from $t = 0$ to the equilibrium. To present our results, we focus on four strategies that play a primordial role in the phenomenon. They are :

- Always Defect (ALLD) : numbers 1, 2, 3 and 4 ;
- Always Cooperate (ALLC) : numbers 20, 24, 28 and 32 ;
- Tit for Tat (TFT) : number 22 ;
- the Retaliator (RET) : number 18.

Figures 3 and 4 represents the evolution of the proportion of these strategies, and of the average score over the population, starting from a uniformly distributed distribution. In both case, and also in all the other experiments that we laid, the scenario leading to cooperation is the same.

At time zero, when the population is uniformly distributed, the malevolent strategies as ALLD realise the best score. They are so the first to grow, at the expense of dumb benevolents as ALLC that quickly disappears. In the uniformly distributed initial population, the susceptible benevolents as RET and TFT do better than ALLC because they avoid being exploited by malevolents, but worse than ALLD because they do not exploit non-susceptible benevolents. So, there quickly happens a situation very different from the initial one, where defection dominates, the population being mainly composed with ALLD, and a little of RET and TFT.

When opposed to itself ALLD realises the poor score of permanent mutual defection, i.e. $L\,P$. When opposed to RET and TFT, it gains a very little more because it exploits the opponent at the first iteration, its score is then $T + (L-1)\,P$. During this confrontation, the benevolent RET or TFT realises the lesser, but close, score $S + (L-1)\,P$. It is when two benevolent are opposed that a significant difference appears. Two benevolent realises the score $L\,R$, that is the best over the population. We verify that this advantage is enough to have benevolent recover the delay that they gain in front of malevolents. So, this is now the susceptible benevolents that grow, and especially the unforgiving RET.

Because they have destroyed their spring of reward constituted by the ALLC population, the malevolents now decrease at great speed. RET continues to grow and achieves a very wide majority of the population, establishing cooperation. A population almost entirely composed with RET is a very stable equilibrium of the system. Other benevolents strategies may sometimes cohabit. Because of the behaviour of RET, everything seems to them as if they were in an ALLC population. From the point of view of malevolents, the population is mainly composed with individual behaving almost as ALLD. So, they realise the poor score of mutual defection and may not survive. RET is still lightly the best, because he avoids the total rout sustained by ALLC when they are opposed to the rare malevolents.

By this mechanism, the combination of benevolence and susceptibility (cf. section 2) of RET guarantees the stability of cooperation. RET and TFT are the only two strategies that are at the same time benevolent and succeptible [8]. Finer studies are needed to understand why TFT does not play a similar role, but it is a fact that simulations always show an advantage in RET intransigence [9]. Some experiments were laid by preventing the creation of RET individuals, reducing so the search space to the other 31 strategies. We saw then that TFT takes the place of RET, although it takes more time to have cooperation emerge. Moreover cooperation does not emerges, if we suppress RET and TFT from the search space. This confirm the primordial importance of benevolent and susceptible strategies in this phenomenon.

All the results presented above are obtained with proportional selection or trivial topology of mutation. In the special case where elitist selection is associated with structured topology, the scenario is almost the same, excepted that instead of an equilibrium, it is a cyclic oscillation between TFT, RET and sometimes other benevolents, that is the attractor of the dynamic.

For the sake of completeness, we conclude this section with the description of the influence of all the parameters (listed in section 5). It is striking to see that most of them have only a weak influence on the behaviour of the system.

---

[8] benevolence traduces by the schemata ($\emptyset \mapsto$ C) and ($R \mapsto$ C), what leaves 8 benevolent strategies (even numbers between 18 and 32). Susceptibility implies ($S \mapsto$ D) and ($P \mapsto$ D) and so, only two strategies may be benevolent and susceptible. TFT's indulgence is the schema ($T \mapsto$ C) replaced by ($T \mapsto$ D) in RET.

[9] a possible explanation is the better score realised by RET in front of strategy 16 that always cooperate excepted at the first iteration.

IPD parameters : We tried two sets of value for (T, R, P, S) that are $(3, 2, 1, 0)$ and the more classical $(5, 3, 1, 0)$, and a very wide variety of values for L (including odd, even and prime numbers), without a significant modification of the observations. However it is possible to change the behaviour of the system with extreme values for these parameters. The combination

$$S \ll P, \qquad P \simeq R, \qquad R \ll T$$

and a short length L (e.g. $(T, R, P, S) = (41, 21, 20, 0)$ and $L = 11$), is unfair with susceptible benevolents and stop the scenario at its first stage, when malevolence dominates (cf.previous discussion about the rise of RET and the fall of ALLC).

Population size S : Until now, we have only described the theoretical behaviour of the model with infinite population. As we explained in section 3, it is obtained by representing the population under the form of a set of real numbers $\{q_i\} \in S_{32}$. Of course, the observed behaviour of the system with finite population tends to it when the population size grows. The size needed to have convergence varies with the model used, GA or ES. With elitist selection, a population of 8 individuals, i.e. 25% of the search space, is enough to have the system evolve accordingly to its theoretical behaviour. That is a quite good result when compared to the proportional selection performances. As a matter of fact, a conformable behaviour of the GA model is not obtained until a population size by 96 individuals, i.e. 300% of the search space. We may understand that the proportional selection, because it is somehow more subtle than the rough Max operator of elitist selection, needs bigger populations, but the observed difference was unexpected and strongly confirms the quote from Bäck and Schwefel (1993) in section 3. However, this very low observed performance of the GA model may be explained by the absence of any recombinative operator as crossover, and thus of a real schema processing that is supposed to be the heart of GA. As we argue in section 7, we believe that the use of crossover may reduce the population size needed, but that it would not change the qualitative behaviour of the system.

Topology of mutation : The results presented in the graphics were all obtained with trivial topology of mutation, i.e. when every strategy may be created from every strategy by mutation. As we explained, the use of structured topology with elitist selection modify the behaviour of the model in the fact that the attractor is not an equilibrium anymore, but it becomes a cycle. With proportional selection, the use of structured topology instead of trivial topology has almost no influence, it just modifies a little the distribution at equilibrium.

Selection operator : The difference between the two alternatives for selection appears clearly on the graphics. The GA'proportional selection is a smooth operator that induces continuity in the behaviour of the model. On the contrary, the ES'elitist selection, because it uses the max operator, has a brutal behaviour leading to discontinuity and non derivability of observable statistics. It is also to be noted that elitist selection realises exact optimisation for a large variety of parameters, but when it does not succeed, it falls an the opposite extreme and realises very low performances.

Temperature θ : As explained in the previous discussion, the temperature is a major parameter that determines all the behaviour of the system.

Initial population : Excepted for very special case, we did not see any influence of the initial distribution of strategies in the population. The attractor of the dynamic, equilibrium or cycle, seems always to be the only existing one.

7 Conclusion

We are now interested in the range of our results. We would like to say that :

> *If the temperature of the genetic melting is not too high and the population size not too low, cooperation naturally emerges under the action of an evolutionary dynamic. Benevolent but susceptible individuals play a major role in this phenomenon and they ensure the stability of cooperation ;*

and so, modify slightly Axelrod's thesis by giving no advantage to indulgence.

The first objection that may be raised is that our system does not represent the generality of evolutionary dynamics, in particular it does not use any recombinative operator as GA'crossover. It seems to us that this critic is well founded with regard to the conclusion that we derived about the algorithm behaviour in general. In particular the observed requirement for a large population size in the GA model (i.e. with proportional selection), may be explained by the absence of crossover and so, of a real schema processing that is supposed to be the heart of a GA. However, we claim that, with regard to the problem of the emergence of cooperation, the use of other genetic operators would not have change the behaviour of the system in a significant way. To support this conjecture, we argue that our system exhibits a surprising robustness, its behaviour not being altered by a change of the selection operator or of the mutation operator. So, we may expect the observed scenario of the emergence of cooperation to be quite general, and valid for every evolutionary dynamics. We are now improving our software by introducing a crossover. We are also putting in place a real implementation of hypothesis NS' (see the end of section 4), that now applies also with finite populations and so replaces hypothesis NS.

The other restriction to the generality of our results is that all our conclusions were derived by restricting ourselves to the small set of strategies that satisfy hypothesis M_1. Although IPD is surely a reliable model for a lot of field of study, its restriction to 32 strategies is only a small game of dubious interest. However, as we explained in section 3, our motivations were to check the validity of the Axelrod 's results with large populations, and it is not more arbitrary to restrict the search space with hypothesis M_1 than with hypothesis M_3. We think that the possibility to cover all the search space, and thus be able to see when the algorithm converge and when it does not, constitute the main interest of our work, that reinforce Axelrod's one.

It is also striking to see how the observed behaviour of the model, i.e. :

1. growth of malevolents and fall of dumb benevolents,
2. raise of susceptible benevolents and fall of malevolents,
3. establishing of cooperation,

is close to the behaviour of the "human" set of strategies of Axelrod's tournaments (Axelrod 1984). It is a pity that no result is provided in (Axelrod 1987), we may wander about the scenario of the raise of cooperation in Axelrod's GA. It seems that the phenomenon that we observe has a certain level of generality, at least when the symmetry of the search space is preserved, as it is the case with M_1 and M_3.

Because of this strong property, IPD may surely be called an "evolution-easy" game. It is tempting to conclude that evolution is a very powerful dynamic that may solve strong cooperation problems. But, if we keep in mind the definition of strong cooperation problems sketched in section 1, i.e. games where Nash equilibriums are strictly dominated by diametrically opposed situations, we are not sure that IPD is a strong cooperation problem. On the contrary, a study that we laid shows us that the situation is not so simple in our 32-alternatives version of IPD. Thus we may wander about the behaviour of evolving populations in strong cooperation problems bigger than the 2-decisions PD. In all cases, it is nice to see that by iterating it, PD becomes an evolution-easy game.

Our theoretical study of IPD and the results of simulations with strong cooperation problems will be the subject of a latter publication. We believe that it is highly improbable that evolution will lead to optimality in a 50-decisions strong cooperation game. If it turns out to be true, no general conclusion about the power of evolutionary dynamics could be derived. Then our future works would be to apply our approach of evolution with large populations, to other not too big paradigmatic games, and find again the nice point of view on evolution that we had with IPD.

References

Aumann R.J. and Hart S. : *Handbook of Game Theory*, North Holland, 1992.

Axelrod R. : Effective choice in the prisoner's dilemma game, *J. Conflict Resolution*, vol. **24**, pp. 3-25, 1980.

Axelrod R. : *The Evolution of Cooperation*, Basic Books, New York, 1984.

Axelrod R. : The evolution of strategies in the iterated prisoner's dilemma, in *Genetic Algorithms and Simulated Annealing*, R.Davis ed., Pitman, London, 1987.

Bäck T. : Self-adaptation in genetic algorithms, *Towards a practice of autonomous systems : Proceedings of the 1st European Conference on Artificial Life*, F. Varela and P. Bourgine eds., MIT press, Cambridge, pp. 263-271, 1992.

Bäck T. and Hoffmeister F. : Extended selection mechanisms in genetic algorithms, *Proceedings of the Fourth International Conference on Genetic Algorithms and their Applications*, R.K. Belew and L.B. Booker eds., Morgan Kaufmann, San Diego, pp. 92-99, 1991.

Bäck T. and Schwefel H.P. : An overview of evolutionary algorithms for parameter optimization, *Evolutionary Computation*, vol. **1**, pp.1-23, 1993.

Baefsky P. and Berge S.E. : Self-sacrifice, cooperation and aggression in women and varying sex-role orientation, *Personality and Social Psychology Bulletin*, vol. **1**, pp. 296-298, 1974.

Bethlehem D.W. : The effect of westernization on cooperative behaviour in Central Africa, *Int. J. Psychology*, vol. **10**, pp. 219-224, 1975.

Boyd R. and Loberbaum J.P.: No pure strategy is evolutionarily stable in the repeated prisoner's dilemma game, *Nature*, vol. **327**, pp. 58-59, 1987.

Feistel R. and Ebeling W. : *Evolution of Complex Systems : self-organisation*, entropy and development, Kluwer Academic Publishers, Dordrecht, 1989.

Fogel D.B. : Evolving behaviours in the iterated prisoner's dilemma, *Evolutionary Computation*, vol. **1**, pp.77-97, 1993.

Gacôgne L. : *Apprentissage génétique d'une extension flou du dilemme itéré du prisonnier*, Rapport n. 94/15, LAFORIA, Université Paris 5, Paris, 1994.

Goldberg D.E. : *Genetic Algorithms in Search, Optimization and Machine Learning*, Addison Wesley, Reading, 1989.

Herdy M. : Application of Evolutionsstrategie to discrete optimization problems, in *Parallel Problem Solving from Nature*, 1st workshop, H.P. Schwefel and R. Männer eds., Springer, pp. 188-192, 1991.

Hoffmeister F. and Bäck T. : Genetic Algorithms and Evolution Strategies : similarities and differences, in *Parallel Problem Solving from Nature*, 1st workshop, H.P. Schwefel and R. Männer eds., Springer, pp. 455-469, 1991.

Holland J. : *Adaptation in natural and artificial systems*, Ann Arbor : the University of Michigan press, Cambridge, 1975.

Lindgren K. and Nordal M.G. : Cooperation and community structure in artificial ecosystem, *Artificial Life*, vol. **1**, pp. 15-37, 1994.

Luce R. and Raiffa H. : *Games and Decisions*, John Wiley and sons, New York, 1957.

Maynard-Smith J. : *Evolution and the Theory of Game, Cambridge University Press*, Cambridge, 1975.

Mühlenbein H. : Darwin's continent cycle theory and its simulation by the Prisoner's Dilemma, in *Towards a practice of autonomous systems : Proceedings of the 1st European Conference on Artificial Life*, P. Bourgine and F. Varela eds., MIT press, Bradford Books, pp. 236-244, 1992.

Riker W. and Brams J.S. : The paradox of vote trading, *American Political Science Review*, vol. **67**, pp. 1235-1247, 1973.

Schelling C.T. : *The Strategy of Conflicts*, Harvard University Press, Harvard, 1960.

Rechenberg I. : *Evolutionsstrategie : Optimierung technischer Systeme nach Prinzipien der biologischen Evolution*, Frommann-Holzboog, Stuttgart, 1973.

Schwefel H.P. : *Numerische Optimierung von Computer-Modellen mittels der Evolutionsstrategie*, Interdisciplinary Systems Research, Birkhäuser, Basel, 1977.

Schwefel H.P. : *Numerical optimization of computer models*, John Wiley and sons, New York, 1981.

Werner G. and Dyer M. : Evolution of communication in artificial organisms, in *Artificial Life 2*, Addison Wesley, pp. 659-687, 1991.

Table 1. The 32 strategies for IPD that satify M_1.

1 : **ALLD** $\emptyset \mapsto$ D $P \mapsto$ D $T \mapsto$ D $S \mapsto$ D $R \mapsto$ D	2 : **ALLD** $\emptyset \mapsto$ D $P \mapsto$ D $T \mapsto$ D $S \mapsto$ D $R \mapsto$ C	3 : **ALLD** $\emptyset \mapsto$ D $P \mapsto$ D $T \mapsto$ D $S \mapsto$ C $R \mapsto$ D	4 : **ALLD** $\emptyset \mapsto$ D $P \mapsto$ D $T \mapsto$ D $S \mapsto$ C $R \mapsto$ C
5 : $\emptyset \mapsto$ D $P \mapsto$ D $T \mapsto$ C $S \mapsto$ D $R \mapsto$ D	6 : $\emptyset \mapsto$ D $P \mapsto$ D $T \mapsto$ C $S \mapsto$ D $R \mapsto$ C	7 : $\emptyset \mapsto$ D $P \mapsto$ D $T \mapsto$ C $S \mapsto$ C $R \mapsto$ D	8 : $\emptyset \mapsto$ D $P \mapsto$ D $T \mapsto$ C $S \mapsto$ C $R \mapsto$ C
9 : $\emptyset \mapsto$ D $P \mapsto$ C $T \mapsto$ D $S \mapsto$ D $R \mapsto$ D	10 : $\emptyset \mapsto$ D $P \mapsto$ C $T \mapsto$ D $S \mapsto$ D $R \mapsto$ C	11 : $\emptyset \mapsto$ D $P \mapsto$ C $T \mapsto$ D $S \mapsto$ C $R \mapsto$ D	12 : $\emptyset \mapsto$ D $P \mapsto$ C $T \mapsto$ D $S \mapsto$ C $R \mapsto$ C
13 : $\emptyset \mapsto$ D $P \mapsto$ C $T \mapsto$ C $S \mapsto$ D $R \mapsto$ D	14 : $\emptyset \mapsto$ D $P \mapsto$ C $T \mapsto$ C $S \mapsto$ D $R \mapsto$ C	15 : $\emptyset \mapsto$ D $P \mapsto$ C $T \mapsto$ C $S \mapsto$ C $R \mapsto$ D	16 : $\emptyset \mapsto$ D $P \mapsto$ C $T \mapsto$ C $S \mapsto$ C $R \mapsto$ C
17 : $\emptyset \mapsto$ C $P \mapsto$ D $T \mapsto$ D $S \mapsto$ D $R \mapsto$ D	18 : **RET** $\emptyset \mapsto$ C $P \mapsto$ D $T \mapsto$ D $S \mapsto$ D $R \mapsto$ C	19 : $\emptyset \mapsto$ C $P \mapsto$ D $T \mapsto$ D $S \mapsto$ C $R \mapsto$ D	20 : **ALLC** $\emptyset \mapsto$ C $P \mapsto$ D $T \mapsto$ D $S \mapsto$ C $R \mapsto$ C
21 : $\emptyset \mapsto$ C $P \mapsto$ D $T \mapsto$ C $S \mapsto$ D $R \mapsto$ D	22 : **TFT** $\emptyset \mapsto$ C $P \mapsto$ D $T \mapsto$ C $S \mapsto$ D $R \mapsto$ C	23 : $\emptyset \mapsto$ C $P \mapsto$ D $T \mapsto$ C $S \mapsto$ C $R \mapsto$ D	24 : **ALLC** $\emptyset \mapsto$ C $P \mapsto$ D $T \mapsto$ C $S \mapsto$ C $R \mapsto$ C
25 : $\emptyset \mapsto$ C $P \mapsto$ C $T \mapsto$ D $S \mapsto$ D $R \mapsto$ D	26 : $\emptyset \mapsto$ C $P \mapsto$ C $T \mapsto$ D $S \mapsto$ D $R \mapsto$ C	27 : $\emptyset \mapsto$ C $P \mapsto$ C $T \mapsto$ D $S \mapsto$ C $R \mapsto$ D	28 : **ALLC** $\emptyset \mapsto$ C $P \mapsto$ C $T \mapsto$ D $S \mapsto$ C $R \mapsto$ C
29 : $\emptyset \mapsto$ C $P \mapsto$ C $T \mapsto$ C $S \mapsto$ D $R \mapsto$ D	30 : $\emptyset \mapsto$ C $P \mapsto$ C $T \mapsto$ C $S \mapsto$ D $R \mapsto$ C	31 : $\emptyset \mapsto$ C $P \mapsto$ C $T \mapsto$ C $S \mapsto$ C $R \mapsto$ D	32 : **ALLC** $\emptyset \mapsto$ C $P \mapsto$ C $T \mapsto$ C $S \mapsto$ C $R \mapsto$ C

Fig. 1. Quality of the equilibrium point as a function of the temperature θ, note that, in the case of elitist selection, the fall from 100% of cooperation to 100% of defection is at $\theta \simeq 2.06$, and the stage at 100% of defection continues over $\theta = 100$ (experience laid with $(T, R, P, S) = (3, 2, 1, 0), L = 30$, infinite population and trivial topology).

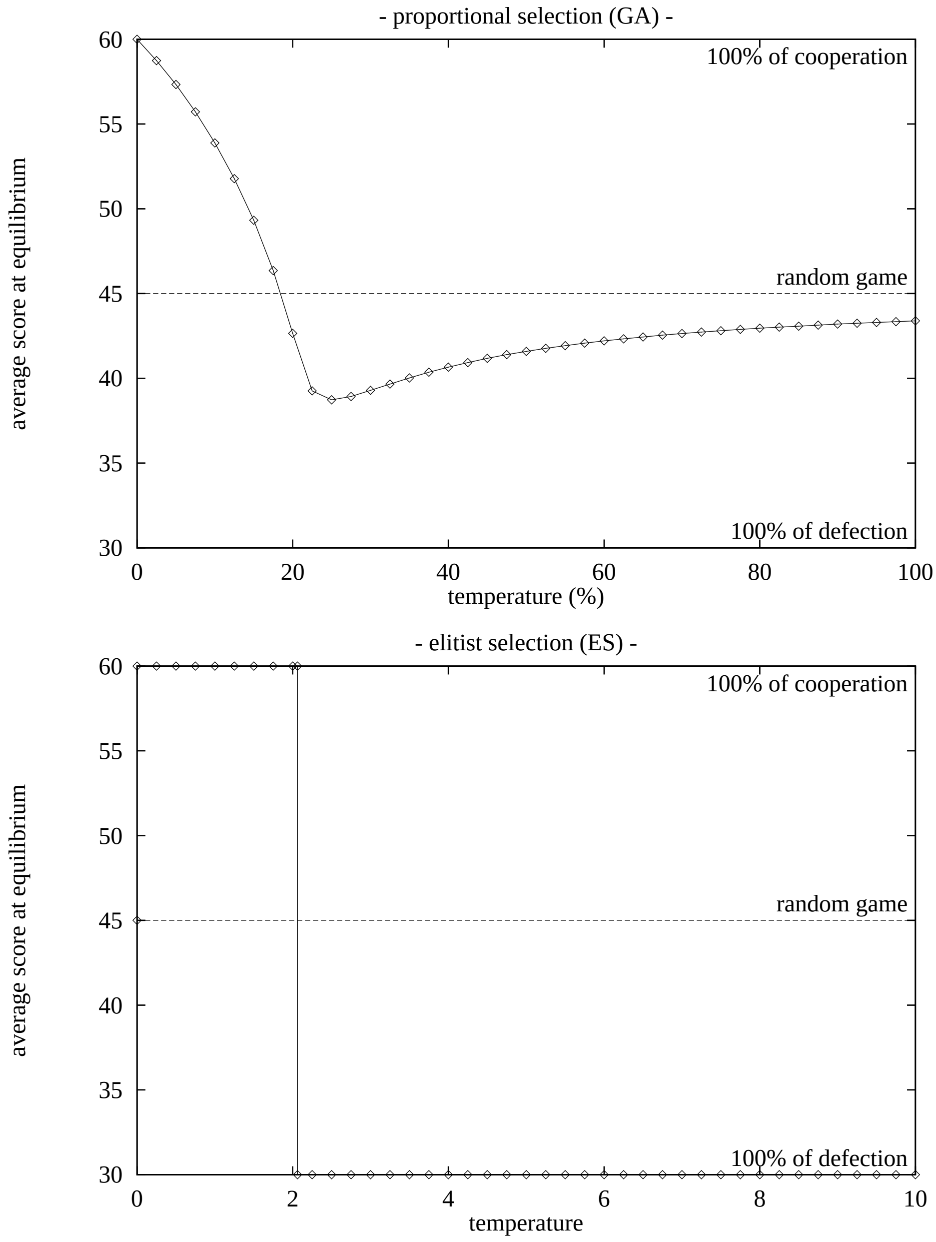

Fig. 2. Population at equilibrium, with proportional selection $\theta = 10\%$ and the average score at equilibrium is 53.89, with elitist selection $\theta = 1$ and the average score at equilibrium is 60 (100% of cooperation) (experience laid with $(T, R, P, S) = (3, 2, 1, 0)$, $L = 30$, infinite population and trivial topology).

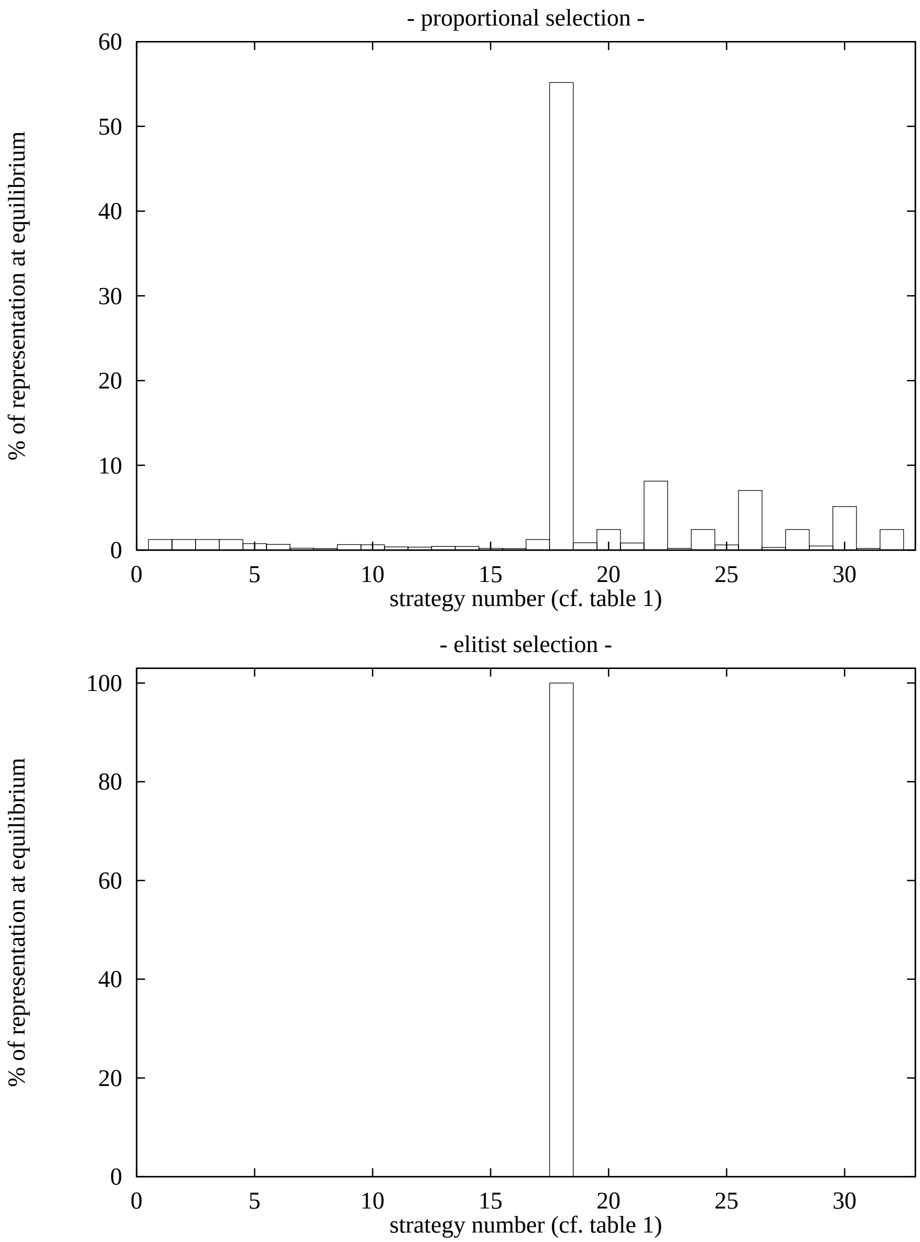

Fig. 3. Evolution of the population through time, starting from a uniformly distributed population and using proportional selection, the average score at equilibrium is 58.76 (experience laid with $(T,R,P,S) = (3,2,1,0), L = 30$, infinite population, $\theta = 2.5\%$ and trivial topology).

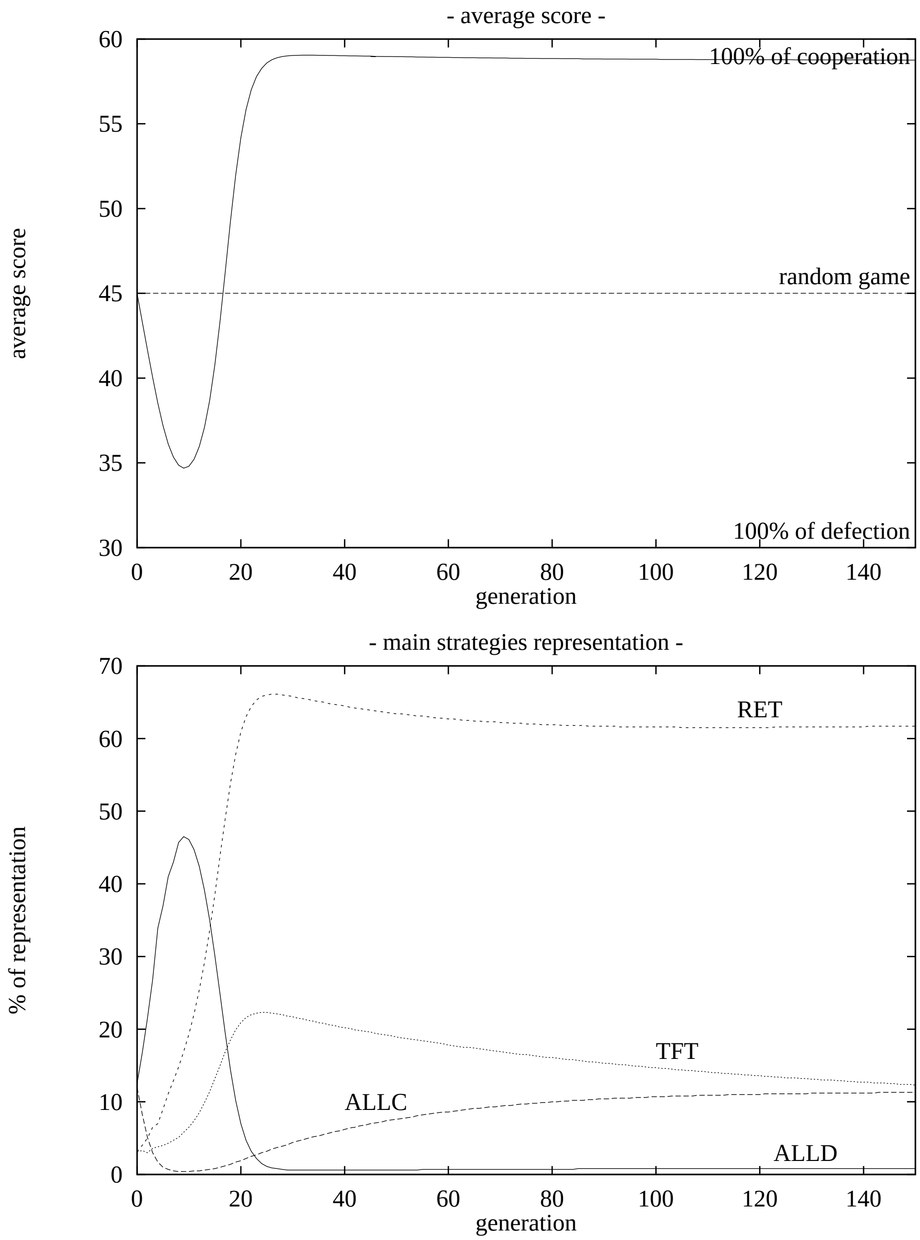

Fig. 4. Evolution of the population through time, starting from a uniformly distributed population and using elitist selection, the average score at equilibrium is 60 (100% of cooperation) (experience laid with $(T, R, P, S) = (3, 2, 1, 0)$, $L = 30$, infinite population, $\theta = 0.75\%$ and trivial topology).